

# Two-Phase Response-Dependent Sampling Designs For Time-to-Event Analysis

Patricia Judd

MSc candidate in Statistics

*Supervisor:* Dr. Yildiz Yilmaz

Department of Mathematics and Statistics, Memorial University

July 13, 2016 @ 11:00 am - 11:59 am  
HH-3017

## **Abstract**

Measuring expensive covariates for all subjects within a cohort may not be a feasible option due to a study's budgetary or logistical constraints. As a result of such limitations, we need to consider sampling designs that account for subjects that have missing data. To design a study allowing incomplete covariate data for some subjects, it is better to employ a cost-efficient sampling design, which balances the efficiency of parameter estimates and power of association tests with the sample size. Response-dependent sampling is a cost-efficient sampling approach. In this approach a subset of subjects is selected from a cohort, based on the response variable (and inexpensive covariates), which has already been gathered for all subjects in the cohort. In our study, we focus on response-dependent two-phase sampling designs. During the first phase of the sampling design, all members in a cohort are measured for the response variable and the inexpensive covariates. In phase 2, a subset of the cohort is selected, based on the response variable obtained in phase 1, and the expensive covariate(s) will be measured only for those selected. In our study, the response variable that determines which individuals are selected for phase 2 is a continuous time-to-event variable; wherein this type of the response variable may be subject to censoring. The most common response-dependent sampling design for time-to-event data is the case-cohort sampling design. We explore variations of case-cohort design which give more efficient association estimates for a given sample size. We stratify cases and/or non-cases based on the observed time-to-event values and apply basic stratified sampling. Different strata sampling proportions change the efficiency of association estimates.